PERFORMANCE & IMPACT EVALUATION (P&IE)

Annual Report, Year 2 (FY2014)

PERFORMANCE AND IMPACT EVALUATION (P&IE)

THE USAID/UGANDA SCHOOL HEALTH AND READING PROGRAM

OCTOBER 31, 2014

PN 7384; USAID Contract N0: AID-617-C-12-00006

PRESENTED TO:

USAID/Uganda

Joseph Mwangi

PRESENTED BY:

NORC at the University of Chicago

4350 East-West Highway, 8th Floor

Bethesda, MD 20814

Telephone: (301) 634-9413

Fax: (301) 634-9301

DISCLAIMER

The author's views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

CONTENTS

A. OVERVIEW

NORC at the University of Chicago (NORC), in partnership with the Panagora Group, is conducting a third party performance and impact evaluation of the School Health and Reading Program (SHRP). SHRP, as implemented by RTI, comprises of two separate activities:

1) Result 1: Improved Early Grade Reading and Transition to English
2) Result 2: Improved HIV/AIDS Knowledge, Attitudes and Practice

RESULT 1: IMPROVED EARLY GRADE READING AND TRANSITION TO ENGLISH

For the Result 1 intervention, SHRP focuses on the nexus of language, pedagogy, and instructional materials to significantly improve students' early grade reading and P3 literacy scores, as well as bring to scale a "Ugandan led 'reading policy'" (RTI International, 2012, p. 1). The Early Grade Reading (EGR) intervention will be implemented at multiple levels.

- At the school level, the intervention provides training to teachers in early grade literacy instruction using students' mother-tongue in P1-P3 and with a transition to English in P4;

- At the district level, instructional and assessment materials are being developed for P1-P4 in the students' mother tongue;

- At the national level, MoES systems and pedagogical and language framework will be strengthened to support mother-tongue based EGR and transition to English. Support will also be provided to strengthen policies related to reading, as well as increase advocacy for reading at multiple levels (e.g. student, teacher, school, district, and national).

Specifically, Under Result 1, SHRP aims to train approximately 12,000 teachers in early grade reading and provide effective reading instruction to almost 1,000,000 learners in primary school levels 1-4 (P1-P4). In collaboration with the National Curriculum Development Centre (NCDC), SHRP will distribute almost 2 million reading primers to 3,300 schools in 30 districts. With funding from the GPE to the MoES, the program will support the MoES to scale up the reading intervention to an additional 1.4 million children in 2,644 primary schools and 56 districts. Thus, by the end of the project, SHRP is expected to improve the instruction and learning environment of 2.8 million learners in almost 6,000 schools and 86 districts, eventually leading to improved literacy skills.

RESULT 2: IMPROVED HIV/AIDS KNOWLEDGE, ATTITUDES, AND PRACTICES

The Result 2 intervention's goal is to improve students' knowledge, attitudes and practices regarding HIV/AIDS. Improving HIV/AIDS education and health supporting attitudes and behaviors will be done by 1) improving MoES planning of the HIV prevention response; 2) improving coordination between MoES and other stakeholders; and 3) supporting the school-level interventions focused on HIV/AIDS and health education for primary grades P4-P7 and grades S1-S4 at the secondary school level. At the school level, to reach a target of 500,000 learners with well-defined messages and activities, SHRP will focus on training teachers at primary and post-primary schools on HIV/AIDS and health topics and in guidance and counseling; developing guidelines for the rollout of the PIASCY package of minimum school interventions (see Box 1); supporting teachers to deliver the enhanced PIASCY program via monitoring and support supervision, referral information, and message support; strengthening the School Family Initiative (SFI) (see Box 1) to reach learners on a co-curricular basis; and increasing and institutionalizing parental and community involvement in school-based HIV/AIDS education.

Since the inception of the P&IE Project in October 2012, the NORC/Panagora team has worked in close collaboration with the RTI team to design and implement a rigorous impact evaluation and a performance evaluation, all within the framework of USAID's Collaboration, Learning, and Adaptation (CLA) Agenda.

We are pleased to submit to USAID/Uganda the Year 2 Annual Report recording progress made on the P&IE project during its second year of implementation.

Box 1: Minimum package of school HIV interventions:

➢ Integrate HIV/AIDS education into classroom subject lessons
➢ Integrate HIV/AIDS issues into co-curricular activities
➢ Provide referrals for HIV and AIDS services
➢ Have participatory and active HIV/AIDS talking compound (HIV/AIDS information on school grounds)
➢ Provide counseling and guidance of pupils by trained HIV/AIDS counselors
➢ Carry out the **School Family Initiative (SFI)**:
"SFI uses small learner groups or "School Families" under the care of a teacher who plays a parental role for the group. The purpose…is to provide a mechanism for in-depth information sharing, guidance, care and support to learners in the area of HIV and AIDS, general health and life skills to small groups of 20-25 pupils…happens on a weekly basis during the school term."

(SHRP Year 2 Work Plan, p. 30)

B. A SNAPSHOT OF YEAR 2 ACTIVITIES

Between October 1, 2013 and October 31, 2014, NORC and Panagora staff:

Completed the first annual impact evaluation analysis for Result 1. P&IE staff completed the first annual impact evaluation of Result 1 activities conducted under SHRP using the baseline and Round 2 EGRA data from Cluster 1 schools in 11 districts collected in February-March and October-November 2013, respectively. The impact evaluation design for the SHRP Result 1 intervention uses both the random assignment of schools to treatment and control groups within SHRP intervention districts (experimental design) and the selection of matched comparison districts (quasi-experimental) in which SHRP is not operating. The experimental design allows us to isolate the effect of school-level interventions from district-level interventions, while the inclusion of non-intervention districts allows us to measure the impact of the district level interventions and the combined district+school level

intervention package. Within this framework, we used ordinary least squares multinomial regressions to evaluate the impact of SHRP on a range of early grade reading outcomes – letter sound knowledge, word segmenting, non-word decoding, oral reading fluency, reading comprehension and English receptive vocabulary - for the four languages covered in Year 1 (Luganda, Ateso, Leblano, and Runyankore/Rukiga,) and English. Different models were used to check the robustness of the results.

In April 2014, we prepared and submitted to USAID the first Impact Evaluation Report for the P&IE Project, documenting the impact of SHRP on different reading outcomes during the first year of its implementation. Box 2, on the following page, presents findings of the impact evaluation.

Box 2: Main Findings of the Year 1 Impact Evaluation:

The impact of the SHRP program on early reading skills has been very modest during the first year of implementation. Specifically:

School-level interventions had:

➢ No impact on any EGRA subtask for the Runyankore/Rukiga, Lango, and Ateso subgroups
➢ Positive impact letter sound knowledge and in the proportion of learners that can identify at least one letter sound in the Luganda subgroup.
➢ Positive impact in the proportion of learners that can identify at least one letter sound in English (although only significant with school fixed effects model)
➢ Positive impact on Luganda word segmenting scores
➢ No impact on higher level literacy skills for the Luganda subgroup or in English for any subgroup

School+district-level interventions had:

➢ No impact on any EGRA subtask for Lango, and Ateso subgroups
➢ No impact on any EGRA subtask in English for any subgroup
➢ Modest positive impact on Luganda letter sound knowledge, word segmenting, nonword decoding, oral reading fluency and reading comprehension
➢ Negative impact on letter sound knowledge and on the proportion of students who could not read any letter sounds in the Runyankore/Rukiga subgroup

---

Conducted data quality assessments (DQA) for three rounds of EGRA. In Year 2 of the P&IE Project, we conducted data quality assessment tasks for three rounds of data collection: the Cluster 1 Round 3 data collection (October 2014); and Cluster 2, baseline (February 2014) and Round 2 (October 2014) data collections. For each round of data collection, the P&IE team reviewed and provided detailed feedback to the implementing partner (IP) on all data collection instruments, both from the perspective of NORC's role as evaluator and data quality reviewer; reviewed and provided feedback on enumerator training manuals; observed enumerator training and pilot test of instruments; and travelled to a selection of districts to observe data collection activities. For the most part, we were able to provide timely feedback to the IP, who took the feedback into consideration, whenever feasible[1].

During Year 2, following each data collection, NORC staff reviewed datasets for quality. These data quality reviews were conducted for the EGRA Cluster 1 Round 2 dataset, which factored into the first impact analysis, along with the Cluster 1 baseline data.

---

[1] Detailed accounts of all DQA activities, including key observations and feedback, are described in Semi-Annual reports and accompanying Data Quality Assessment Reports, submitted to USAID in April and October 2014.

Completed a Midterm Performance Evaluation. P&IE staff conducted a Mid-Term Performance Evaluation and subsequently prepared a detailed "Mid-Term Performance Evaluation Report" for SHRP, which was submitted to USAID on September 30, 2014.

The SHRP Mid-Term Evaluation was organized around six evaluation questions covering design, implementation, results, and sustainability, effective use of funds, and management and learning; a summary of the main findings, conclusions, and recommendations for each question are presented in the Midterm Evaluation report. The evaluation utilized formative, cross-sectional, descriptive, and analytical methods and drew from the comprehensive monitoring data and observations that the NORC team has been collecting since the inception of the project (described below); data and information gathered through reviews of SHRP reports, work plans, PMPs, and other documents; and qualitative data obtained through an intensive round of key informant interviews, focus group discussions, and classroom observations conducted at the national, district, and school level, during a two week trip to Uganda by the Performance Evaluation team. We employed purposeful sampling of schools and districts. Data quality and analysis was validated through triangulation of multiple sources and stakeholders.

This performance evaluation complements P&IE's impact evaluation; while the impact evaluation measures program impact and the degree to which the end results – in particular, reading outcomes -are being achieved, the performance evaluation provided rich qualitative data on stakeholder perceptions, beliefs, and thinking, which has been analyzed by a cross-disciplinary team of health, education, and development experts to develop findings, conclusions, and recommendations.

Given the breadth and scope of the performance evaluation and its findings, we have opted not to present a snapshot of findings in this Annual Report. A comprehensive set of findings, organized by the six evaluation questions for both Results and 1 and 2, are available in the Midterm Performance Evaluation Report.

Systematically observed and documented implementation activities as input to the performance evaluation. Local P&IE staff continued to observe key SHRP events/activities throughout the year, including trainings of trainers and teachers, language board meetings, leadership and management training for district and school-level education officials, EGRA data collection in schools for Clusters and 1 and 2; and accompanied SHRP team members on support supervision field visits to monitor implementation of Result 1 activities. Per our established practice, observers recorded both positive and constructive observations in tools developed by P&IE. These observations served as critical input for the mid-term SHRP performance evaluation; they also form the basis for appreciative and constructive feedback to the IP in monthly CLA meeting between SHRP and P&IE team members. In addition, these observation activities, coupled with regular communication with SHRP team leaders have proven invaluable in keeping the P&IE team abreast of implementation progress, delays in the rollout of activities, and bottlenecks/problems in the field that have implications for the impact evaluation.

Operationalized USAID Collaboration, Learning and Adaptation (CLA) priorities through regular feedback discussions with the IP. To implement the CLA component of the P&IE contract, we provided performance feedback, drawn from the observation activities and tools discussed above, to the

SHRP team on a monthly basis. The purpose of the monthly performance feedback session is to provide both appreciative and constructive feedback particularly on elements of performance where real-time feedback will engender learning and help to strengthen performance, ultimately leading to better outcomes. The performance feedback is drawn directly from documented observations of events and activities, which are synthesized into feedback memos submitted to SHRP leadership prior to the monthly meetings. During Year 2 the P&IE and SHRP team staff engaged in seven monthly feedback sessions.

The SHRP team has found the third party perspective on performance informative and useful. The P&IE team has obtained useful clarifications and insights into the implementation activities during the feedback sessions. We are pleased that this somewhat sensitive exercise is perceived as a valued and welcome opportunity to achieve its underlying purpose of improving performance and results achievement in real time.

C. RISKS, CHALLENGES, AND MODIFICATIONS TO THE IMPACT EVALUATION

During the past 12 months, several challenges/risks to the impact evaluation designs have emerged. Some of these risks have been resolved or mitigated, while others still remain to be addressed. We have described these challenges at great length in the April 2014 and October 2014 P&IE Semi-Annual Reports. Below, we present a summary of the challenges.

- During the Cluster 2 EGRA training and pilot test, the P&IE team observers noted some issues related to the implementation of three specific EGRA subtasks – Letter Sound Knowledge and Word Segmenting, and Oral Passage Reading - and that could have negative implications for the impact evaluation. Specifically, SHRP was using very stringent requirements for accepting letter sounds as correct; for example, while the EGRA toolkit states that "For consonants that can represent more than one sound (i.e., c, g), either answer is acceptable. For vowels, either the short or long sound is accepted (/i/ as in pin or as in pine)," in the SHRP implementation of EGRA only one sound per vowel was being accepted as correct. As well, local pronunciations of words – e.g. "muzzah" for mother – were being marked as incorrect. This raises the concern that learners who actually know correct letter sounds are assessed as not knowing them, since trainers were instructed during training to mark as wrong any very slight deviation from the "ideal" sound of a letter. This approach of "teaching to the test" can bias the assessment in favor of treatment schools, where students are being taught one correct letter sound or a specific pronunciation of a word, relative to control schools, where a broader set of letter sounds and pronunciations are being taught.

- Possible contamination of Result 1 control group from two sources.

    — Because the SHRP team, for various reasons, is not planning to expand SHRP implementation to new districts for Cluster 1, they are implementing Result 1 activities in control CCTs in the 11 original districts starting in 2014 in order to meet target numbers of trained teachers. However, they plan to exclude the control schools within the control CCTs that were selected for the EGRA data collection and intervene only in the schools

from control CCTs which have not been included in the EGRA data collection. The SHRP M&E Team Lead has assured us that no teachers in any grade (P1 through P4) in the EGRA control schools will be trained; nor will instructional materials be distributed to these schools. CCTs associated with these control clusters will be strictly instructed not to provide any assistance to these control schools. This strict exclusion of control schools from treatment is critical for the integrity of the impact evaluation design. While SHRP staff has assured us that no control schools will receive the Result 1 interventions, we are nonetheless concerned by the possibility of contamination through CCTs or spillover of materials. Any contamination of the control schools will lead to underestimation of the effects of the SHRP Result 1 interventions.

— In October 2014, we learned from RTI that Mango Tree Project was working in Otuke, a control district for the SHRP evaluation, and providing their literacy intervention to two control schools in the SHRP sample. This occurred despite careful coordination between Mango Tree and SHRP. The schools in question were replaced for others, however the replacement schools have no baseline and their usefulness is limited. In addition, it is possible that some contamination has already occurred as we do not have any type of control over Mango Tree activities in the district.

▪ Non-systematic replacement of sample schools. During the Cluster 2 baseline data collection in Mbale district, the SHRP team opted to exclude control schools that use or were presumed to use Luganda and English instead of Lumasaaba as the medium of instruction. However, appropriate procedures to replace these schools (following the replacement rule provided) were not followed. Two of these non-Lumasaaba instruction schools were replaced by schools in which the medium of instruction is Lumasaaba; these replacements were picked from the list of preselected schools designated as replacements. The rest of the non-Lumasaaba instruction schools in the district sample were neither assessed nor replaced. We indicated to the IP and USAID that this approach was neither appropriate to keeping the integrity of a random sample nor conducive to comparing SHRP schools to the average public school in Uganda. First, replacing sample schools with hand-picked replacements creates problems with the sample balance. Second, the aim of the evaluation is to assess reading ability of learners in English and local language. While it is not possible to test them in the local language (Lumasaaba, in this case) in schools that do not teach in Lumasaaba, it would still have been possible to test student's performance in English. As such, NORC's Evaluation Expert urged SHRP staff to conduct the EGRA in English in these schools as soon as we learned of the situation. However, the SHRP team did not comply with this request in a timely manner. Therefore, NORC decided that the impact analysis will need to exclude Mbale district altogether.

▪ Manafwa district encountered a serious crisis created by teacher transfers in the region. We learned during field observations that most of the teachers trained by SHRP in January 2014 in this region have been transferred to other schools: four of the treatment schools visited by our local staff did not have a trained P1 teacher, because s/he had been transferred. It will be critical to have information about the whereabouts of teachers trained by SHRP, since transfers of trained teachers away from treatment schools will have a severe effect on the impact evaluation. If these teachers end up at control schools, the impacts will be even more skewed. We will work with the IP and through our performance evaluation to try and capture the movement of trained teachers between schools.

- Sample size adjustments between rounds. The IP has changed sample sizes of each cohort between rounds. For the first cohort of students (Cluster 1), the February 2013 baseline included 280 schools to allow for analysis of 3 treatment arms, controls in treatment districts, and controls in comparison districts. However, the Cluster 1 Round 2 data collection conducted in October 2013, RTI collected data only in a subsample of treatment schools (168 of the 280), since a decision was made to only focus on one (and not 3) treatment. For Cluster 1, Round 3 in October 2014, however, RTI reverted back to data collection from 280 schools to account for that fact that the 3 treatment arms were implemented in the second year. This use of unbalanced panels does not preclude us from conducting a rigorous evaluation; however, it makes the process less transparent and prevents us from having measurements year by year without loss of information and precision. A similar change happened for Cluster 2. In this case, RTI requested NORC to calculate a sample size large enough to be able to analyze results at the district level. At baseline, in February 2014, data was collected from enough number of schools to calculate impact at district level; however, in for the first follow up in October 2014, the IP decided that district level analysis was not of interest and reduced the data collection to a subsample of the original schools. Although we will not be able to say anything about impact at district level, if properly implemented, this change should not prevent us from analyzing results at language level. In general, NORC recommends following the original samples over time to produce a more streamlined and transparent process and more comparable results across year.

D. LESSONS LEARNED DURING YEAR 2

The ongoing observation of project implementation and discussion of implementation progress and performance, as well as problems and bottlenecks observed, with the implementing partner served three critical functions:

(1) It has afforded us an opportunity to convey concerns to the IP in real time, thereby allowing SHRP staff to make implementation adjustments and midstream course correction as appropriate.

(2) It resulted in an incredibly rich and comprehensive midterm evaluation that differed greatly from most performance evaluations conducted by USAID partners. In lieu of limiting the PE to interviews and observation conducted at a distinct point in time during the life of the project, the P&IE team was able to draw on information gathered over a two-year period from observation of SHRP implementation activities on an ongoing and continuous basis, recording implementation progress, deviations from implementation plan, and problems and bottlenecks encountered in the course of project implementation. These observations laid the foundations for an in-depth performance evaluation that took into consideration implementation realities throughout the entire life of SHRP and informed the point-in-time key informant interviews and focus group discussions that were conducted by the P&IE team in July 2014.

(3) It has been critical for identifying possible threats and risks to the impact evaluation, allowing us to make adjustments to the sampling and evaluation designs, as needed, to mitigate these risks. Even when such adjustments are not possible, understanding the implementation realities and deviations from plan that affect the impact evaluation is important in the analysis and results interpretation phase.

E. YEAR 3 WORKPLAN

Below we present a detailed work plan schedule for activities for the period November 1, 2014 – October 31, 2015.

Work Plan	
Activities	*Month*
Incorporate comments and revisions to Deliverable 5: Receive comments from USAID on Midterm Performance Evaluation report, Semi-Annual Report #4 (with DQR), Annual Report #2, and Year 2 Work Plan; make necessary changes and submit revised deliverables.	**October 31 – November 30, 2014**
Conduct Dissemination Workshops for Midterm PE and First Annual IE results The P&IE Team will organize and implement a dissemination workshop in Uganda to share findings from the performance evaluation and first annual impact evaluation.	**Dec 2014**
Complete ex-post data quality assessments for EGRA, Cluster 1, Round 3 and Cluster 2, Round 2 data collection NORC will review raw and clean datasets provided by the SHRP M&E team for quality, consistency, coverage, and usefulness for the impact evaluation NORC will provide direct feedback to the SHRP implementation and M&E team on any observed issues/concerns; we will document all quality control activities and findings/observations, and mitigating actions in the Data Quality Assurance Report that will be submitted to USAID on April 30, 2015, along with Semi-Annual Report #5.	**December 1, 2014 – January 31, 2015**
Support sample selection and data collection, including data quality assessments, for EGRA, Cluster 3 baseline data collection and Round 2, Cluster 1 Round 4, and Cluster 2 Round 3; and KAP midline — Support selection of schools for evaluation sample, ensuring that treatment and control school selection adheres to the sampling plan — Review and provide feedback on all data collection instruments, ensuring that the data being collected link back to evaluation questions/impact indicators, and that the instruments are of high quality — Review and provide feedback on training manuals, data collection protocols/plans, and quality control procedures for field work, and tablet software being used for data collection — Participate, for quality assurance purposes, in enumerator training and pilot testing of instruments – NORC's Senior Literacy Expert and Survey Specialist will travel to Uganda to participate in the separate trainings for the Results 1 and Results 2 data collections	**January – October 2014** EGRA, Cluster 3, baseline: Feb 2015 EGRA, Cluster 3, baseline: Feb 2015: Oct-Nov 2015

Work Plan	
Activities	*Month*
— Conduct field observations during the first two weeks of data collection – NORC's Survey Specialist, Resident Evaluation Manager, and Senior HIV/AIDS Specialist will travel to the field to observe field work for both Results 1 and Results 2 data collections — Conduct additional field observations in later stages of data collection – to be undertaken by NORC's Resident Evaluation Manager, and Senior HIV/AIDS Specialist — Conduct quality reviews of data, as it is uploaded onto RTI's servers from tablets; we expect to be able to conduct real-time data quality reviews if RTI provides NORC with access to the server to which data is being uploaded daily NORC will provide direct feedback to the SHRP implementation and M&E team on any observed issues/concerns; additionally, all quality control activities and findings/observations, and mitigating actions will be documented in the Data Quality Assurance Reports that will be submitted to USAID on April 30 and October 31, 2015.	EGRA, Cluster 1, Round 4 and Cluster 2, Round 3: Oct-Nov 2015 KAP midline: Oct 2015 NORC experts' travel will be timed to coincide with training and pilot testing for these data collection activities
Monitor Implementation of SHRP and Evaluation Design NORC's Resident Evaluation Manager and Senior HIV/AIDS Evaluator will make regular visits to intervention schools to observe the implementation of Result 1 and Result 2 activities as a means of gathering information for the performance evaluation, such that the midterm and final performance evaluations will not be informed by information collected only at two points in time, but instead will be fed organically by information gathered in real time, throughout the project. These field visits and regular meetings with SHRP Results Teams will also serve to ensure that program implementation is adhering to the impact evaluation designs. Where implementation deviates from plan, NORC will discuss with USAID and RTI, and discuss adjustments to implementation plan and/or evaluation design. NORC will engage in ongoing dialogue and discussion with RTI about evaluation implementation, and challenges. Semi-Annual and Annual Reports will include details of evaluation implementation, problems encountered, midstream course correction/modification to design and findings during the reporting period.	**Nov 2014 – October 2015**
Continue monthly feedback sessions and quarterly performance review meetings with SHRP Team. During these monthly meetings, the P&IE team, in keeping with the CLA Agenda, will synthesize and share with the SHRP leadership observations and	**Monthly, Nov 2014 – October 2015**

Work Plan	
Activities	*Month*
recommendations based on observation and monitoring activities conducted by P&IE local staff. Each session will have a corresponding feedback report.	
Conduct data analysis and prepare Second Annual Impact Evaluation Report Conduct the second annual impact analysis of EGRA Cluster 1 and Cluster 2 schools, the multiple rounds of data collected in 2013 and 2014 and include this analysis in the First Impact Evaluation Report, due on April 30, 2015. This report will not contain an impact analysis for Result 2, since the second round of KAP data will not be collected until October 2015. Impact analysis for Result 2 activities will only be possible after the midline KAP data collection in October 2015.	**Dec 2014 – April 2015**
Develop Final Performance Evaluation Implementation Plan The PE Team, led by Panagora, will work in close consultation with the Evaluation Expert in developing the detailed Final Performance Evaluation Plan, due in October 2015.	**Sep-Oct 2015**

Deliverable Schedule: Year 3	
Deliverable	*Due Date*
Second Impact Evaluation Report Semi-Annual Report #5, with DQR of Oct 2014 data collection	Apr 30, 2015
Semi-Annual Report #6, with DQR, Annual Report #3 Year 4 Work Plan with Final Performance Evaluation Implementation Plan	Oct 31, 2015